9

# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/027,817 | 12/21/2001 | Hamid R. Berenji | IIS-110 | 9496 |

7590 12/31/2003

Andrew V. Smith
Sierra Patent Group, Ltd.
P.O. Box 6149
Stateline, NV 89449

| EXAMINER | |
|---|---|
| LIU, JOSHUA C | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2121 | 5 |

DATE MAILED: 12/31/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

2

| Office Action Summary | Applicati n No. 10/027,817 | Applicant(s) BERENJI ET AL. | |
|---|---|---|---|
| | Examiner Joshua C Liu | Art Unit 2121 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *12/21/201 (eff. filing date 3/30/2001)*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-12* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-12* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *21 December 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. §§ 119 and 120**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

13)☒ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

a) ☐ The translation of the foreign language provisional application has been received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1:78.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

## DETAILED ACTION

1. Claims 1-12 have been examined.

### *Specification*

2. Applicant is reminded of the proper language and format for an abstract of the disclosure.

The abstract should be in narrative form and generally limited to a single paragraph on a separate sheet within the range of 50 to 150 words. It is important that the abstract not exceed 150 words in length since the space provided for the abstract on the computer tape used by the printer is limited. The form and legal phraseology often used in patent claims, such as "means" and "said," should be avoided. The abstract should describe the disclosure sufficiently to assist readers in deciding whether there is a need for consulting the full patent text for details.

The language should be clear and concise and should not repeat information given in the title. It should avoid using phrases which can be implied, such as, "The disclosure concerns," "The disclosure defined by this invention," "The disclosure describes," etc.

3. The abstract of the disclosure is objected to because the Applicant recites "said at least one" on L. 11 of the abstract. Correction is required. See MPEP § 608.01(b).

4. The disclosure is objected to because of the following informalities:

- In the Incorporation by Reference section on Pg. 11-12, there are several recitations of "emph" that appear to be typographical errors.

### *Claim Objections*

5. Claims 4-5 are objected to because of the following informalities:

- Claim 4 recites "software program for providing" on L. 1, which may be non-statutory. The Examiner suggests the following correction: "computer-readable medium having".

- Claim 5 recites "comprising the operations" on L. 2-3, which is unclear. The Examiner suggests the following correction: "comprising the steps of". Appropriate correction is required.

***Claim Rejections - 35 USC § 101***

1. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

2. Claims 1-3 and 5-7 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. Claims 1-3 and 5-7 are not explicitly or implicitly declared to produce tangible and concrete results. On that basis alone, those claims are clearly non-statutory.

Regardless of whether the claims are in the technological arts, none of them are limited to practical applications in the technological arts. Examiner finds that *In re Warmerdam*, 33 F.3d 1354, 31 USPQ2d 1754 (Fed. Cir. 1994), controls the 35 U.S.C. §101 issues on that point for reasons made clear by the Federal Circuit in *AT&T Corp. v. Excel Communications, Inc.*, 50 USPQ2d 1447 (Fed. Cir. 1999). Specifically, the Federal Circuit held that the act of:

> "taking several abstract ideas and manipulating them together adds nothing to the basic equation." *Excel* at 1453 (quoting *In re Warmerdam*, 33 F.3d 1354, 1360 (Fed. Cir. 1994)).

Examiner finds that Applicant's "fuzzy-logic rules", "reinforcement learning algorithm", "optimum value", and "update equation" are just such abstract idea.

Examiner bases his position upon guidance provided by the *Excel* court on *Warmerdam*, as interpreted by the court's finding in *Excel*. This set of precedents is

within the same line of cases as the *Alappat-State Street Bank* decisions and is in complete agreement with those decisions. *Warmerdam* is consistent with *State Street*'s holding that:

> "we hold that *the transformation of data, representing* ***discrete dollar amounts****, by a machine through a series of mathematical calculations into a final share price*, constitutes a practical application of a mathematical algorithm, formula, or calculation because it produces a 'useful, concrete and tangible result.' – *a final share price momentarily fixed for recording purposes and even accepted and relied upon by regulatory authorities and in subsequent trades.*" *State Street Bank* at 1601.

True enough, that case later eliminated the "business method exception" in order to show that business methods were not per se non-statutory, but the court clearly *did not* go so far as to make business methods *per se statutory.* A plain reading of the excerpt above shows that the court was *very specific* in its definition of the new *practical application.* It would have been much easier for the court to say that "business methods were per se statutory" than it was to define the practical application in the case as "the transformation of data, **representing discrete dollar amounts**, by a machine through a series of mathematical calculations into a final share price..."

The court was being very specific.

Additionally, the court was also careful to specify that the "useful, concrete and tangible result" it found was "a final share price momentarily fixed for recording purposes and even accepted and **relied upon** **by regulatory authorities and in subsequent trades.**"

Applicant cites no such specific results to define a useful, concrete and tangible result. Neither does Applicant specify the associated practical application with the kind of specificity the Federal Circuit used.

Furthermore, the *Warmerdam* court held that:

"th disp sitiv issu for assessing compliance with Section 101 in this case is whether the claim is for a process that goes beyond simply **manipulating 'abstract ideas' or 'natural phenomena'**... As the Supreme Court has made clear, '[a]n idea of itself is not patentable, ... *taking several abstract ideas and manipulating them together adds nothing to the basic equation*." *Warmerdam*, 31 USPQ2d at 1759 (emphasis added).

In the present case, the Examiner finds that Applicant manipulated a set of abstract "fuzzy-logic rules", "reinforcement learning algorithm", "optimum value", and "update equation" to solve mathematical problems in the **abstract**. Under *Warmerdam*, the result of such manipulations is not statutory.

Since *Warmerdam* is within the *Alappat-State Street Bank* line of cases, it takes the same view of "useful, concrete, and tangible" requirement the Federal Circuit applied in *State Street Bank.* Therefore, under *State Street Bank*, this could not be a "useful, concrete and tangible result." There is only manipulation of abstract ideas.

The Federal Circuit validated the use of *Warmerdam* in its more recent *Excel* decision. The court noted that:

"Finally, the decision in *In re Warmerdam*, 33 F.2d 1354, 31 USPQ2d 1754 (Fed. Cir. 1994) is not to the contrary. *** The court found that the claimed process did nothing more than **manipulate basic mathematical constructs** and concluded that '*taking several abstract ideas and manipulating them together adds nothing to the basic equation*'; hence, the court held that the claims were properly rejected under §101 ... Whether one agrees with the court's conclusion on the facts, the holding of the case is a straightforward application of the basic principle that mere laws of nature, natural phenomena, and abstract ideas are not within the categories of inventions or discoveries that may be patented under §101." *Excel Communications, Inc.*, 50 USPQ2d 1447, 1453 (Fed. Cir. 1999).

The fact that the invention is merely the manipulation of *abstract ideas* is indisputable. The object referred to by Applicant's abstract word "fuzzy-logic rules", "reinforcement learning algorithm", "optimum value", and "update equation" are simply mathematical/logical constructs in the abstract. Consequently, the necessary conclusion under *AT&T*, *State Street*, and *Warmerdam*, is straightforward and clear. The claims take several abstract ideas (i.e. linear and non-linear correction of error

signal) and manipulate them together adds nothing to the basic equation. Claims 1-3 and 5-7 are rejected under 35 U.S.C. 101.

- Claim 1 is directed to a software program, the software program comprising a database of fuzzy logic rules and a reinforcement learning algorithm, which reveal no tangible subject matter as described in Detailed Description of the Preferred Embodiment. The Examiner suggests that the Applicant correct this error by replacing "A software program for providing" with "A computer-readable medium having".
- Claims 2-3 appear to recite additional software elements without any tangible subject matter. The Examiner suggests that the Applicant correct this error by replacing "A software program for providing" with "A computer-readable medium having".
- Claim 5 is directed to a method of controlling a system. However, the method comprises the steps of mapping input data to output commands and updating the fuzzy logic rules based on effects on the system state, which reveal no tangible subject matter as described in Detailed Description of the Preferred Embodiment.
- Claims 6-7 appear to recite additional steps that do not affect any tangible device.

### ***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the

invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

7. Claims 1-3, 5-7, and 9-11 are rejected under 35 U.S.C. 103(a) as being unpatentable over Yamakawa et al (US Patent Number 6,633,858; Filed 10/14/2003) in view of Konda et al (IDS Reference 27; Published 11/1999).

**Claim 1**

Claim 1 recites

A software program for providing instructions to a processor which controls a system for applying actor-critic based fuzzy reinforcement learning, comprising:

(a) a database of fuzzy-logic rules for mapping input data to output commands for mapping a system state; and

(b) a reinforcement learning algorithm for updating the fuzzy-logic rules database based on effects on the system state of the output commands mapped from the input data, and

(c) wherein the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations.

- Regarding claim 1, Yamakawa discloses a computer-readable medium having instructions to a processor which controls a system for applying actor-critic based fuzzy reinforcement learning (Yamakawa Fig. 19; Col 15 L. 27-35, "A control program… read a control program."; Col 2 L. 1-7, "An actor-critic model… an actor module."; Col 2 L. 36-40, "The present invention… to be improved."; Fig. 3), comprising: (a) a database of fuzzy-logic rules for mapping input data to output commands for mapping a system state (Yamakawa Fig. 17; Col 11 L. 18-28, "The movable state… movable state list."; Col 12 L. 13-17, "The selector… becomes the shortest."; Col 14 L. 41-45, "A landmark database… landmark position function."), and (b) a reinforcement learning algorithm for updating the fuzzy-logic rules database based on the effects on the system state (Yamakawa Fig. 16; Col 11 L. 18-53, "The movable state model… the learning process.").

However, Yamakawa does not explicitly teach that (c) the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations. Konda teaches a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations (Applicant's Background of the Invention, Pg. 7 L. 6-8, "Recently, Konda and Tsitsiklis... approximation techniques."; Konda §4, "The best that one... becomes small (infinitely often)."), which –selects every action with a non-zero probability and still converge for continuous state-action spaces (Applicant's Background of the Invention, Pg. 7 L. 12-13, "They also suggested... assumptions are satisfied."), and applies to high-dimensional problems and is mathematically sound (Konda §5, "our algorithm apply... certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the art to modify Yamakawa, in view of Konda, by using a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations.

**Claim 2**

Claim 2 recites "The software program of Claim 1, wherein the reinforcement learning algorithm is based on an update equation including a derivative with respect to

said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered."

- Regarding claim 2, see §103 rejection for claim 1, and (Applicant's Background of the Invention, Pg. 8 Eq. 11-12; Konda §2, "In reference to Assumption (A1), note... = $\nabla \ln \mu_\theta(x, u)$."), which –converges if the learning rate sequences satisfy certain conditions (Applicant's Background of the Invention, Pg. 9, "The above algorithm converges... for either $\alpha_t$ or $\beta_t$.") and applies to high-dimensional problems (Konda §5, "our algorithm apply... certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the art to further modify Yamakawa, in view of Konda, by basing the reinforcement learning algorithm on an update equation including a derivative with respect to said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered.

**Claim 3**

Claim 3 recites "The software program of Claim 2, wherein the reinforcement learning algorithm is configured to update the at least one parameter based on said update equation."

- Regarding claim 3, see §103 rejection for claim 2.

**Claim 5**

Claim 5 recites

A method of controlling a system including a processor for applying actor-critic based fuzzy reinforcement learning, comprising the operations:

(a) mapping input data to output commands for modifying a system state according to fuzzy-logic rules;

(b) updating the fuzzy-logic rules based on effects on the system state of the output commands mapped from the input data; and

(c) converging at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations.

- Regarding claim 5, Yamakawa discloses a method of controlling a system including a processor for applying actor-critic based fuzzy reinforcement learning (Yamakawa Fig. 19; Col 15 L. 27-35, "A control program… read a control program."; Col 2 L. 1-7, "An actor-critic model… an actor module."; Col 2 L. 36-40, "The present invention… to be improved."; Fig. 3), comprising the steps of: (a) mapping input data to output commands for modifying a system state according to fuzzy-logic rules (Yamakawa Fig. 17; Col 11 L. 18-28, "The movable state… movable state list."; Col 12 L. 13-17, "The selector… becomes the shortest."; Col 14 L. 41-45, "A landmark database… landmark position function."), and (b) updating the fuzzy-logic rules database based on the effects on the system state (Yamakawa Fig. 16; Col 11 L. 18-53, "The movable state model… the learning process.").

However, Yamakawa does not explicitly teach (c) converging at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations. Konda teaches a method for converging at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations (Applicant's Background of the Invention, Pg. 7 L. 6-8, "Recently, Konda and Tsitsiklis… approximation techniques."; Konda §4, "The best that one… becomes small (infinitely often)."), which –selects every action with a non-

zero probability and still converge for continuous state-action spaces (Applicant's Background of the Invention, Pg. 7 L. 12-13, "They also suggested… assumptions are satisfied."), and applies to high-dimensional problems and is mathematically sound (Konda §5, "our algorithm apply… certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the art to modify Yamakawa, in view of Konda, by using a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations.

**Claim 6**

Claim 6 recites "The method of Claim 5, wherein the updating operation includes taking a derivative with respect to said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered."

- Regarding claim 6, see §103 rejection for claim 5, and (Applicant's Background of the Invention, Pg. 8 Eq. 11-12; Konda §2, "In reference to Assumption (A1), note… = $\nabla \ln \mu_\theta(x, u)$."), which –converges if the learning rate sequences satisfy certain conditions (Applicant's Background of the Invention, Pg. 9, "The above algorithm converges… for either $\alpha_t$ or $\beta_t$.") and applies to high-dimensional problems (Konda §5, "our algorithm apply… certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the art to further modify Yamakawa, in view of Konda, by having the step of updating include taking a derivative with respect to said at least one parameter of a logarithm of a

probability function for taking a selected action when a selected state is encountered.

**Claim 7**

Claim 7 recites “The method of Claim 6, wherein the updating operation includes updating the at least one parameter based on said derivative.”

- Regarding claim 7, see §103 rejection for claim 6.

**Claim 9**

Claim 9 recites

A system controlled by an actor-critic based fuzzy reinforcement learning algorithm which provides instructions to a processor of the system for applying actor-critic based fuzzy reinforcement learning, comprising:
(a) the processor;
(b) at least one system component whose actions are controlled by said processor;
(c) at least one storage medium accessible by said processor, including data stored therein corresponding to:
(i) a database of fuzzy-logic rules for mapping input data to output commands for modifying a system state; and
(ii) a reinforcement learning algorithm for updating the fuzzy-logic rules database based on effects on the system state of the output commands mapped from the input data, and
(iii) wherein the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations.

- Regarding claim 9, Yamakawa discloses a system controlled by an actor-critic based fuzzy reinforcement learning algorithm which provides instructions to a process of the system for applying actor-critic based fuzzy reinforcement learning (Yamakawa Fig. 19; Col 15 L. 27-35, “A control program… read a control program.”; Col 2 L. 1-7, “An actor-critic model… an actor module.”; Col 2 L. 36-40, “The present invention… to be improved.”; Fig. 3), comprising: (a) the processor (Yamakawa Fig. 19); (b) at least one system component controlled by said processor (Yamakawa Fig. 13 Element 10); (c) at least one storage medium

(Yamakawa Fig. 19), including data stored therein corresponding to: (i) a database of fuzzy-logic rules for mapping input data to output commands for mapping a system state (Yamakawa Fig. 17; Col 11 L. 18-28, “The movable state… movable state list.”; Col 12 L. 13-17, “The selector… becomes the shortest.”; Col 14 L. 41-45, “A landmark database… landmark position function.”), and (ii) a reinforcement learning algorithm for updating the fuzzy-logic rules database based on the effects on the system state (Yamakawa Fig. 16; Col 11 L. 18-53, “The movable state model… the learning process.”).

However, Yamakawa does not explicitly teach that (iii) the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations. Konda teaches a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations (Applicant’s Background of the Invention, Pg. 7 L. 6-8, “Recently, Konda and Tsitsiklis… approximation techniques.”; Konda §4, “The best that one… becomes small (infinitely often).”), which –selects every action with a non-zero probability and still converge for continuous state-action spaces (Applicant’s Background of the Invention, Pg. 7 L. 12-13, “They also suggested… assumptions are satisfied.”), and applies to high-dimensional problems and is mathematically sound (Konda §5, “our algorithm apply… certain convergence properties.”). Therefore, it would have been obvious to one of ordinary skill in the

art to modify Yamakawa, in view of Konda, by using a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations.

**Claim 10**

Claim 10 recites "The system of Claim 9, wherein the reinforcement learning algorithm is based on an update equation including a derivative with respect to said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered."

- Regarding claim 10, see §103 rejection for claim 9, and (Applicant's Background of the Invention, Pg. 8 Eq. 11-12; Konda §2, "In reference to Assumption (A1), note... = $\nabla$ ln $\mu_\theta(x, u)$."), which –converges if the learning rate sequences satisfy certain conditions (Applicant's Background of the Invention, Pg. 9, "The above algorithm converges... for either $\alpha_t$ or $\beta_t$.") and applies to high-dimensional problems (Konda §5, "our algorithm apply... certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the art to further modify Yamakawa, in view of Konda, by basing the reinforcement learning algorithm on an update equation including a derivative with respect to said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered.

**Claim 11**

Claim 11 recites "The system of Claim 10, wherein the reinforcement learning algorithm is configured to update the at least one parameter based on said update equation."

- Regarding claim 11, see §103 rejection for claim 10.

8. Claims 4, 8, and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sutton et al (IDS Reference 25; Published 1998) in view of Yamakawa et al (US Patent Number 6,633,858; Filed 10/14/2003), and further in view of Konda et al (IDS Reference 27; Published 11/1999).

**Claim 4**

Claim 4 recites "The software program of any of Claims 1-3, wherein the system includes a wireless transmitter."

- Regarding claim 4, Sutton discloses that a mobile telephone system, which inherently comprises wireless transmitters, blocks calls less frequently when using reinforcement learning method to allocate channels compared to using "fixed assignment" method or "borrowing with directional channel locking" method (Sutton Fig. 11.10; Pg. 282-283, "Singh and Bertsekas… for large systems.").

  However, Sutton does not teach which reinforcement learning method is used and how the reinforcement learning method is implemented. Yamakawa discloses a computer-readable medium having instructions to a processor which controls a system for applying actor-critic based fuzzy reinforcement learning (Yamakawa Fig. 19; Col 15 L. 27-35, "A control program… read a control program."; Col 2 L. 1-7, "An actor-critic model… an actor module."; Col 2 L. 36-

40, “The present invention… to be improved.”; Fig. 3), comprising: (a) a database of fuzzy-logic rules for mapping input data to output commands for mapping a system state (Yamakawa Fig. 17; Col 11 L. 18-28, “The movable state… movable state list.”; Col 12 L. 13-17, “The selector… becomes the shortest.”; Col 14 L. 41-45, “A landmark database… landmark position function.”), and (b) a reinforcement learning algorithm for updating the fuzzy-logic rules database based on the effects on the system state (Yamakawa Fig. 16; Col 11 L. 18-53, “The movable state model… the learning process.”), --which provides a problem solver that allows the calculation cost in executing an action to be reduced and the flexibility against a change of a goal state to be improved (Yamakawa Col 2 L. 36-40, “The present invention… to be improved.”).

However, Yamakawa does not explicitly teach that (c) the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations. Konda teaches a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations (Applicant's Background of the Invention, Pg. 7 L. 6-8, “Recently, Konda and Tsitsiklis… approximation techniques.”; Konda §4, “The best that one… becomes small (infinitely often).”), which –selects every action with a non-zero probability and still converge for continuous state-action spaces (Applicant's Background of the Invention, Pg. 7 L. 12-13, “They also suggested…

assumptions are satisfied."), and apply to high-dimensional problems and are mathematically sound (Konda §5, "our algorithm apply... certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the art to modify Yamakawa, in view of Konda, by using a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations.

**Claim 8**

Claim 8 recites "The method of any of Claims 5-7, wherein the system includes a wireless 5 transmitter."

- Regarding claim 8, Sutton discloses that a mobile telephone system, which inherently comprises wireless transmitters, blocks calls less frequently when using reinforcement learning method to allocate channels compared to using "fixed assignment" method or "borrowing with directional channel locking" method (Sutton Fig. 11.10; Pg. 282-283, "Singh and Bertsekas... for large systems.").

  However, Sutton does not teach which reinforcement learning method is used and how the reinforcement learning method is implemented. Yamakawa discloses a method of controlling a system including a processor for applying actor-critic based fuzzy reinforcement learning (Yamakawa Fig. 19; Col 15 L. 27-35, "A control program... read a control program."; Col 2 L. 1-7, "An actor-critic model... an actor module."; Col 2 L. 36-40, "The present invention... to be improved."; Fig. 3), comprising the steps of: (a) mapping input data to output

commands for modifying a system state according to fuzzy-logic rules (Yamakawa Fig. 17; Col 11 L. 18-28, "The movable state… movable state list."; Col 12 L. 13-17, "The selector… becomes the shortest."; Col 14 L. 41-45, "A landmark database… landmark position function."), and (b) updating the fuzzy-logic rules database based on the effects on the system state (Yamakawa Fig. 16; Col 11 L. 18-53, "The movable state model… the learning process."), --which allows the calculation cost in executing an action to be reduced and improves the flexibility against a change of a goal state (Yamakawa Col 2 L. 36-40, "The present invention… to be improved.").

However, Yamakawa does not explicitly teach that (c) converging at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations. Konda teaches a method for converging at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations (Applicant's Background of the Invention, Pg. 7 L. 6-8, "Recently, Konda and Tsitsiklis… approximation techniques."; Konda §4, "The best that one… becomes small (infinitely often)."), which –selects every action with a non-zero probability and still converge for continuous state-action spaces (Applicant's Background of the Invention, Pg. 7 L. 12-13, "They also suggested… assumptions are satisfied."), and applies to high-dimensional problems and is mathematically sound (Konda §5, "our algorithm apply… certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the

art to modify Yamakawa, in view of Konda, by using a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations.

**Claim 12**

Claim 12 recites "The system of any of Claims 9-11, wherein said at least one system component comprises a wireless transmitter."

- Regarding claim 12, Sutton discloses that a mobile telephone system, which inherently comprises wireless transmitters, blocks calls less frequently when using reinforcement learning method to allocate channels compared to using "fixed assignment" method or "borrowing with directional channel locking" method (Sutton Fig. 11.10; Pg. 282-283, "Singh and Bertsekas… for large systems.").

  However, Sutton does not teach which reinforcement learning method is used and how the reinforcement learning method is implemented. Yamakawa discloses a controlled by an actor-critic based fuzzy reinforcement learning algorithm which provides instructions to a process of the system for applying actor-critic based fuzzy reinforcement learning (Yamakawa Fig. 19; Col 15 L. 27-35, "A control program… read a control program."; Col 2 L. 1-7, "An actor-critic model… an actor module."; Col 2 L. 36-40, "The present invention… to be improved."; Fig. 3), comprising: (a) the processor (Yamakawa Fig. 19); (b) at least one system component controlled by said processor (Yamakawa Fig. 13 Element 10); (c) at least one storage medium (Yamakawa Fig. 19), including data

stored therein corresponding to: (i) a database of fuzzy-logic rules for mapping input data to output commands for mapping a system state (Yamakawa Fig. 17; Col 11 L. 18-28, "The movable state… movable state list."; Col 12 L. 13-17, "The selector… becomes the shortest."; Col 14 L. 41-45, "A landmark database… landmark position function."), and (ii) a reinforcement learning algorithm for updating the fuzzy-logic rules database based on the effects on the system state (Yamakawa Fig. 16; Col 11 L. 18-53, "The movable state model… the learning process."), --which provides a problem solver that allows the calculation cost in executing an action to be reduced and the flexibility against a change of a goal state to be improved (Yamakawa Col 2 L. 36-40, "The present invention… to be improved.").

However, Yamakawa does not explicitly teach that (iii) the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations. Konda teaches a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations (Applicant's Background of the Invention, Pg. 7 L. 6-8, "Recently, Konda and Tsitsiklis… approximation techniques."; Konda §4, "The best that one… becomes small (infinitely often)."), which –selects every action with a non-zero probability and still converge for continuous state-action spaces (Applicant's Background of the Invention, Pg. 7 L. 12-13, "They also suggested…

assumptions are satisfied."), and applies to high-dimensional problems and is mathematically sound (Konda §5, "our algorithm apply… certain convergence properties."). Therefore, it would have been obvious to one of ordinary skill in the art to modify Yamakawa, in view of Konda, by using a reinforcement learning algorithm configured to converge at least one parameter of the system state towards approximately an optimum value following multiple mapping and updating iterations.

***Conclusion***

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Joshua C Liu whose telephone number is (703) 305-6435. The examiner can normally be reached on Monday-Friday, 8:30am-5:15pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Anil Khatri can be reached on (703) 305-0282. The fax phone number for the organization where this application or proceeding is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-3900.

jl

ANIL KHATRI
SUPERVISORY PATENT EXAMINER